

APPROXIMATELY CERTIFYING THE RESTRICTED ISOMETRY PROPERTY IS HARD

JONATHAN WEED

ABSTRACT. A matrix is said to possess the Restricted Isometry Property (RIP) if it acts as an approximate isometry when restricted to sparse vectors. Previous work has shown it to be NP-hard to determine whether a matrix possesses this property, but only in a narrow range of parameters. In this work, we show that it is NP-hard to make this determination for *any* accuracy parameter, even when we restrict ourselves to instances which are either RIP or far from being RIP. This result implies that it is NP-hard to approximate the range of parameters for which a matrix possesses the Restricted Isometry Property with accuracy better than some constant. Ours is the first work to prove such a claim without any additional assumptions.

1. INTRODUCTION

The field of compressed sensing, inaugurated by the seminal paper of Candès and Tao [7], offers an attractive and powerful set of techniques for reconstructing sparse data on the basis of very few measurements. Implementing compressed sensing techniques in practice requires taking measurements according to a matrix with special properties. The most widely known such property is the *restricted isometry property* [6], which requires that the matrix act as an approximate isometry when restricted to sparse vectors.

Definition 1. A matrix $X \in \mathbb{R}^{n \times p}$ possesses the (k, δ) -*restricted isometry property* if it satisfies

$$(1) \quad (1 - \delta)\|u\| \leq \|Xu\| \leq (1 + \delta)\|u\|$$

for all k -sparse vectors $u \in \mathbb{R}^p$. In this case, we write $X \in \text{RIP}(k, \delta)$.

Candès showed [6] that if $X \in \text{RIP}(k, \delta)$ for $\delta < \sqrt{2} - 1$, then an ℓ_1 minimization procedure used with the matrix X exactly recovers k -sparse vectors.

Finding RIP matrices in the most interesting range of parameters is an object of active study. Constructing such matrices deterministically is a hard problem [3, 5, 8], but there are several very simple random methods known to generate RIP matrices with high probability [4, 11]. The fact that these randomized algorithms have a small probability of failure motivates the question of certifying whether a given matrix is RIP:

Problem 1. Given a matrix X , a positive integer k , and $\delta > 0$, is $X \in \text{RIP}(k, \delta)$?

Date: April 4, 2017.

This work was supported in part by NSF Graduate Research Fellowship DGE-1122374. The author would like to thank Philippe Rigollet for helpful discussions.

While previous work has shown Problem 1 to be computationally hard, none of these works have been able to show that Problem 1 is NP-hard in an interesting range of parameters. (See Section 1.4 for a fuller account.) In particular, earlier work has only been able to show the NP-hardness of deciding whether $X \in \text{RIP}(k, \delta)$ for $\delta = 1 - o(1)$. By contrast, the correct question in practice is to decide whether $X \in \text{RIP}(k, \delta)$ for some *constant* δ .

An even more reasonable question in practice is to determine the *approximate* range of parameters for which a matrix possesses the restricted isometry property. We therefore propose the following modification of Problem 1:

Problem 2. Fix constants $\delta \in (0, 1)$ and $\lambda_1, \lambda_2 > 1$. Is $X \in \text{RIP}(k, \delta)$ or is X far from being in $\text{RIP}(k, \delta)$, in the sense that X is not even in $\text{RIP}(k/\lambda_1, \lambda_2\delta)$?

Note that Problem 2 is *easier* than Problem 1, since all that we seek is a procedure to distinguish between two very different situations. For this reason, Problem 2 is known as a *gap problem* in the computational complexity literature. More details about problems of this type appear in Section 1.1

In this work, we show that Problem 2 is NP-hard for all $\delta \in (0, 1)$. This immediately establishes the NP-hardness of Problem 1 as well, and in a much more robust sense than implies by previous work.

We also consider the following two related problems.

Problem 3. Given a matrix X and $\delta > 0$, find the largest positive integer k such that $X \in \text{RIP}(k, \delta)$.

Problem 4. Given a matrix X and positive integer k , find the smallest $\delta > 0$ such that $X \in \text{RIP}(k, \delta)$.

Our results show that Problems 3 and 4 are hard to solve even approximately.

1.1. Gap hardness. Gap problems are part of a broader class of *promise problems*, where the input is guaranteed to fall into one of two classes. In a gap problem, these two classes are assumed to be well separated. Concretely, given a general optimization problem of the form

$$(2) \quad \max_{x \in \mathcal{X}} f(x),$$

and a threshold t , the c -gap problem is to distinguish between

$$\max_{x \in \mathcal{X}} f(x) \leq t \text{ and } \max_{x \in \mathcal{X}} f(x) > ct$$

for some constant $c > 1$. If this gap problem is NP-hard, then it is clearly NP-hard to find a c -multiplicative approximation to (2). For this reason, gap hardness results immediately imply that the approximation problem is also hard [19].

To establish the NP-hardness of Problem 2, we produce a reduction from an NP-hard gap problem, via a reduction that preserves the gap between classes.

We reduce from a problem known as “max positive 1-in-3 SAT.”

Definition 2. Given boolean variables x_1, \dots, x_k , the predicate “exactly one” (or E1) is given by

$$\text{E1}(x_1, \dots, x_k) = \begin{cases} T & \text{if exactly one of the variables} \\ & x_1, \dots, x_k \text{ is true,} \\ F & \text{otherwise.} \end{cases}$$

A *max positive 1-in-3 SAT* instance is a set of E1 clauses $\{c_1, \dots, c_m\}$ each containing at most 3 variables from a set $\{x_1, \dots, x_n\}$. The word “positive” is used to stress that all variables appear in positive form. If $\text{E1}(c_i) = T$, we say that clause i is *satisfied*. In this work, we will consider instances with a further restriction.

Definition 3. A instance of max positive 1-in-3 SAT is called 6-bounded if each variable appears in at most 6 clauses.

Given an instance ϕ of 6-bounded max positive 1-in-3 SAT and an assignment $\mathbf{x} \in \{T, F\}^n$, denote by $\text{val}(\phi, \mathbf{x})$ the proportion of clauses satisfied by the assignment \mathbf{x} . Finally, define

$$\text{val}(\phi) := \max_{\mathbf{x} \in \{T, F\}^n} \text{val}(\phi, \mathbf{x}).$$

If $\text{val}(\phi) = 1$, that is, if there is an assignment satisfying all clauses, we say that ϕ is *satisfiable*.

Our reduction is based on the following proposition.

Proposition 1. *There exists a constant α such that, given a 6-bounded max positive 1-in-3 SAT instance ϕ , it is NP-hard to distinguish between $\text{val}(\phi) = 1$ and $\text{val}(\phi) < (1 - \alpha)$. Moreover, the instances ϕ under consideration can be restricted to contain exactly $9/13$ as many clauses as variables.*

A proof of Proposition 1 appears in Section 2. It is well known that deciding whether an instance ϕ is *satisfiable* is an NP-complete problem [16], and the gap hardness of max positive 1-in-3 SAT (without the 6-boundedness condition) is proved in [12]. Hardness problems of this type were first officially stated in [15], and their NP-hardness follows from the celebrated PCP Theorem [1].

1.2. Main result. We show the following gap hardness result for Problem 2.

Theorem 1. *For all $\delta \in (0, 1)$ there exist constants $\lambda_1, \lambda_2 > 1$ such that, given a matrix X and sparsity parameter k , it is NP-hard to decide whether $X \in \text{RIP}(k, \delta)$ or $X \notin \text{RIP}(k/\lambda_1, \lambda_2\delta)$. Moreover, the claim holds even when $\|Xu\| \leq \|u\|$ for all u .*

For $\lambda_1, \lambda_2 > 1$, the condition that $X \in \text{RIP}(k/\lambda_1, \lambda_2\delta)$ is weaker than $X \in \text{RIP}(k, \delta)$, since the bounds in Equation (1) are weaker and the sparsity condition is stronger, so that Equation (1) is required to hold for a smaller set of vectors. Theorem 1 says that even if X satisfies the strong condition $X \in \text{RIP}(k, \delta)$, it is hard to even certify that it satisfies the weak condition $X \in \text{RIP}(k/\lambda_1, \lambda_2\delta)$. The restriction to X such that $\|Xu\| \leq \|u\|$ implies that Problem 2 is hard even when only the *lower* bound of (1) is in question.

Theorem 1 implies the following hardness of approximation results.

Corollary 1. *For all $\delta \in (0, 1)$, there exists a constant $\lambda_1 > 1$ such that it is NP-hard to solve Problem 3 to within a λ_1 factor.*

Corollary 2. *There exists a constant $\lambda_2 > 1$ such that it is NP-hard to solve Problem 4 to within a λ_2 factor.*

1.3. Proof strategy. Suppose we have an instance ϕ of 6-bounded positive 1-in-3 SAT with n variables and m clauses.

Given such an instance, define a $m \times n$ matrix Φ :

$$(3) \quad \Phi_{ij} = \begin{cases} 1 & \text{if variable } j \text{ appears in clause } i, \\ 0 & \text{otherwise.} \end{cases}$$

Any vector $v \in \{0, 1\}^n$ can be interpreted as an assignment of true and false to n variables, where $v_j = 1$ if variable j is true, and $v_j = 0$ otherwise. The definition of Φ implies

$$\begin{aligned} (\Phi v)_i = 1 &\iff \text{clause } i \text{ contains exactly one true variable} \\ &\iff \text{clause } i \text{ is satisfied.} \end{aligned}$$

We obtain that ϕ is satisfiable if and only if there exists a 0-1 vector v such that $\Phi v = \mathbf{1}$, the all-ones vector. On the other hand, if $\text{val}(\phi) \leq 1 - \alpha$, then for all $v \in \{0, 1\}^n$,

$$\|\Phi v - \mathbf{1}\|^2 \geq \alpha m.$$

This implies that being able to solve the optimization problem

$$(4) \quad \min_{v \in \{0, 1\}^n} \|\Phi v - \mathbf{1}\|^2$$

would immediately yield a procedure to compute $\text{val}(\phi)$. Therefore, under the assumption that computing $\text{val}(\phi)$ is intractable, we obtain that the problem in (4) must also be hard to solve. Moreover, the gap hardness of computing $\text{val}(\phi)$ implies that even finding a constant-factor approximation to (4) is NP-hard.

We will construct a matrix X and sparsity parameter k such that solving

$$(5) \quad \min_{u: \|u\|=1, \|u\|_0 \leq k} \|Xu\|$$

is approximately equivalent to solving (4).

The matrix X we construct will contain a rescaled version of Φ as a submatrix. The remaining entries of X will be chosen in such a way to ensure that the sparse vectors u for which $\|Xu\|^2$ is minimized are approximately 0-1 vectors, and hence correspond approximately to feasible vectors v in (4). Then, we will argue that for 0-1 vectors, the values of (4) and (5) are equal up to an additive shift. By carefully controlling the errors at every step, we show knowledge of the value of (5) up to some constant level of accuracy would imply the ability to solve (4), and hence the ability to estimate $\text{val}(\phi)$. The hardness of the later program then completes the proof.

1.4. Prior work. Several papers have shown Problem 1 to be computationally intractable under a number of different assumptions [2, 13, 14, 18].

In [13], the authors analyze a problem similar to our Problem 2; however, they rely on a stronger assumption than $P \neq NP$, and are only able to show that is hard to distinguish between $X \in \text{RIP}(k, \delta)$ and $X \notin \text{RIP}(k', \delta')$ with $|\delta - \delta'|$ approaching zero as the size of the instance increases.

The first two papers to prove the NP-hardness of Problem 1 [2, 18] both rely on the fact that given a matrix $X \in \mathbb{R}^{n \times p}$ and a sparsity parameter k , it is NP-hard to certify whether the kernel of X contains a nonzero k -sparse vector. When no such vector exists, the best that can be said is that

$$\|Xu\| \geq 2^{-\text{poly}(n,p)} \|u\| \quad \text{for all } k\text{-sparse } u \neq 0.$$

These reductions therefore show that certifying $X \in \text{RIP}(k, \delta)$ is hard, but only when $\delta = 1 - \varepsilon$ for some ε that is exponentially small in n and p .

The work most similar to ours is [14], in which the authors raised the same objections we do about the restrictiveness of Problem 1. They also prove that Problem 2 is hard, but only under the *small-set expansion hypothesis* (see [14] for a definition), which asserts that a particular graph problem is NP-hard to approximate. By contrast, we are able to prove directly the NP-hardness of Problem 2 without any additional assumptions.

1.5. Notation and Terminology. Given a vector $x \in \mathbb{R}^d$, we write $\|x\|$ for the ℓ_2 -norm of x , and $\|x\|_0$ for the ℓ_0 -“norm” defined by

$$\|x\|_0 = |\{i \in [d] : x_i \neq 0\}|.$$

If $\|x\|_0 \leq k$, we say that x is k -sparse.

The symbol $\mathbf{1}$ denotes the all-ones vector.

2. BOUNDED MAX POSITIVE 1-IN-3 SAT

In this Section, we prove Proposition 1. We reduce from a problem called max 3SAT-5. An instance of max 3SAT-5 is a CNF formula where each clause contains exactly 3 variables (in positive or negative form) and each variable appears in exactly 5 clauses. The max 3SAT-5 problem is known to be gap hard:

Proposition 2 (Feige [9]). *There exists a constant α' such that, given an instance ψ of max 3SAT-5, it is NP-hard to distinguish between $\text{val}(\psi) = 1$ and $\text{val}(\psi) < (1 - \alpha')$.*

Given an instance ψ of max 3SAT-5 with n variables and $m = \frac{5n}{3}$ clauses, we produce an instance ϕ of 6-bounded max positive 1-in-3 SAT with $n' = 2n + 4m$ variables and $m' = 3m + n = \frac{9}{13}n'$ clauses such that:

- $\text{val}(\psi) = 1 \implies \text{val}(\phi) = 1$,
- $\text{val}(\psi) < (1 - \alpha') \implies \text{val}(\phi) < (1 - \alpha)$, where $\alpha = \alpha'/18$.

The claimed NP-hardness of distinguishing $\text{val}(\phi) = 1$ and $\text{val}(\phi) < (1 - \alpha)$ then follows.

Proof of Proposition 1. We will produce the instance ϕ from ψ in several stages, first by transforming ψ into an instance ψ' of 1-in-3 SAT that contains negated variables, and then transforming ψ' into an instance of *positive* 1-in-3 SAT.

We first produce an instance of 1-in-3 SAT that contains negated variables. Consider a clause $(a \vee b \vee c) \in \psi$, where a, b, c represent arbitrary literals (positive or negative variables). Replace this clause by the three 1-in-3 SAT clauses:

$$\text{E1}(\bar{a}, z_1, z_2), \text{E1}(b, z_2, z_3), \text{E1}(\bar{c}, z_3, z_4),$$

where \bar{x} denotes the negated version of the literal x and z_1, \dots, z_4 are four fresh variables appearing in these three clauses and no others. If $(a \vee b \vee c)$ is satisfied, then there is a setting of z_1, \dots, z_4 to satisfy all three of the new clauses. If $(a \vee b \vee c)$ is not satisfied, then any setting of z_1, \dots, z_4 leaves at least one clause unsatisfied. Repeating this replacement for each clause in ψ yields ψ' .

To obtain a positive instance, replace each occurrence of x_i or \bar{x}_i by the new variable w_i or y_i , respectively, and add the clause

$$\text{E1}(w_i, y_i).$$

Call the resulting positive 1-in-3 SAT instance ϕ . Note that ϕ has $m' = 3m + n$ clauses, and that each variable appears in at most 6 clauses. The instance ϕ involves $n' = 2n + 4m$ variables, of which the $2n$ variables $w_1, \dots, w_n, y_1, \dots, y_n$ correspond to positive and negative versions of $\{x_1, \dots, x_n\}$ in ψ . In particular, we note that $m' = \frac{9}{13}n'$.

If ψ is satisfiable, then clearly ϕ is as well. So suppose that $\text{val}(\psi) < (1 - \alpha')$. Any assignment to the variables in ϕ induces a partial assignment to the variables in ψ in the following way: if only one of the variables w_i or y_i is true, then say x_i is true if w_i is true and x_i is false if y_i is true. If w_i and y_i are both true or both false, then say x_i is indeterminate.

Suppose an assignment to the variables in ϕ leaves b variables in ψ indeterminate. These b variables are involved in at most $5b$ clauses in ψ . Any full assignment to ψ fails to satisfy at least $\alpha'm$ clauses. By removing clauses containing indeterminate variables, it follows that there are at least $(\alpha'm - 5b)_+$ unsatisfied clauses in ψ all of whose variables are determined, where $(a)_+ = \max\{0, a\}$. These $(\alpha'm - 5b)_+$ unsatisfied clauses in ψ correspond to at least $(\alpha'm - 5b)_+$ unsatisfied clauses in ψ' . Moreover, each indeterminate variable corresponds to one additional unsatisfied clause in ϕ , since if x_i is indeterminate then $E1(w_i, y_i) = F$. So the assignment fails to satisfy at least

$$(\alpha'm - 5b)_+ + b \geq \alpha'm/5$$

clauses in ϕ .

Hence

$$\text{val}(\phi) < \frac{m' - \alpha'm/5}{m'} = (1 - \alpha'/18) = (1 - \alpha),$$

as desired. \square

3. PROOF OF MAIN THEOREM

In this Section, we prove Theorem 1, which we recall below.

Theorem 1. *For all $\delta \in (0, 1)$ there exist constants $\lambda_1, \lambda_2 > 1$ such that, given a matrix X and sparsity parameter k , it is NP-hard to decide whether $X \in \text{RIP}(k, \delta)$ or $X \notin \text{RIP}(k/\lambda_1, \lambda_2\delta)$. Moreover, the claim holds even when $\|Xu\| \leq \|u\|$ or all u .*

We first prove the statement for a specific choice of δ , and later show how the proof can be extended to all $\delta \in (0, 1)$.

3.1. Proof overview. The reduction is from 6-bounded positive 1-in-3 SAT. By Proposition 1, there exists a constant α such that it is NP-hard to distinguish satisfiable 6-bounded positive 1-in-3 SAT instances from instances in which only a $1 - \alpha$ fraction of clauses are satisfiable.

Given ϕ with n variables and m clauses, we construct a matrix $\tilde{X} \in \mathbb{R}^{(4n+m) \times 3n}$ with the following three properties:

- (1) If $\text{val}(\phi) = 1$, then there exists a unit vector u with $\|u\|_0 = 2n$ such that

$$\|\tilde{X}u\|^2 = 1/2.$$

- (2) If $\text{val}(\phi) \leq 1 - \alpha$, then there exists a constant $c_1 > 1$ such that every unit vector u with $\|u\|_0 \leq 2c_1n$ satisfies

$$\|\tilde{X}u\|^2 \geq (1 + c_2)/2$$

for some constant $c_2 > 0$.

(3) The matrix \tilde{X} has operator norm at most c_3 for some constant $c_3 > c_2$.

Given \tilde{X} with the above three properties, consider the matrix $X = \frac{1}{c_3}\tilde{X}$. By Property 3,

$$\|Xu\|^2 \leq \|u\|^2$$

for all vectors u . Choose $\delta = 1 - \frac{1+c_2}{2c_3}$ and $\lambda_2 = \frac{2c_3}{2c_3-c_2}$. We obtain the following: if $\text{val}(\phi) < 1 - \alpha$, then for all u such that $\|u\|_0 \leq 2c_1n$

$$\|Xu\|^2 \geq \frac{1+c_2}{2c_3}\|u\|^2 = (1-\delta)\|u\|^2.$$

Conversely, if $\text{val}(\phi) = 1$ then there exists a u satisfying $\|u\|_0 = 2n$ such that

$$\|Xu\|^2 = \frac{1}{2c_3}\|u\|^2 < (1-\lambda_2\delta)\|u\|^2$$

In other words, letting $k = 2c_1n$ and $\lambda_1 = c_1$ yields

$$\begin{aligned} \text{val}(\phi) < 1 - \alpha &\implies X \in \text{RIP}(k, \delta) \\ \text{val}(\phi) = 1 &\implies X \notin \text{RIP}(k/\lambda_1, \lambda_2\delta). \end{aligned}$$

Since it is NP-hard to distinguish between $\text{val}(\phi) = 1$ and $\text{val}(\phi) < 1 - \alpha$, it is also NP-hard to distinguish between $X \in \text{RIP}(k, \delta)$ and $X \notin \text{RIP}(k/\lambda_1, \lambda_2\delta)$.

The matrix \tilde{X} is defined in Section 3.2. We first verify Properties 1 and 3, and then in Section 3.3 reduce the verification of Property 2 to verifying two conditions on the minimizer of the program given in (7). We verify these conditions in Section 3.4. Finally, we show how to extend the proof to general δ in Section 3.5.

3.2. Definition of \tilde{X} . Let ε and ξ be small constants to be chosen later, with $\xi \ll \varepsilon < 1$. Let I be the $n \times n$ identity matrix and $\mathbf{1}$ the all-ones vector of length n , and define

$$P = I - \frac{1}{n}\mathbf{1}\mathbf{1}^\top.$$

P is an orthogonal projection onto the subspace orthogonal to $\mathbf{1}$.

Let $\tilde{X} \in \mathbb{R}^{(4n+m) \times 3n}$ be the following matrix:

$$(6) \quad \tilde{X} = \left(\begin{array}{c|c|c} I & 0 & 0 \\ \hline 0 & I & 0 \\ \hline 0 & 0 & \xi^{-1}P \\ \hline \xi^{-1}I & \xi^{-1}I & -\xi^{-1}I \\ \hline \varepsilon\Phi & 0 & -\varepsilon I' \end{array} \right),$$

where I' is the $n \times n$ identity matrix truncated to have only m rows.

Call a vector $u \in \mathbb{R}^{3n}$ an *assignment vector* if

$$\begin{aligned} \{u_i, u_{i+n}\} &= \{0, 1\} & \text{for } 1 \leq i \leq n, \\ u_j &= 1 & \text{for } 2n < j \leq 3n. \end{aligned}$$

We can interpret such vectors as true-false assignments to n variables by setting $x_i = T$ if $u_i = 1$ and $x_i = F$ if $u_{i+n} = 1$.

Proposition 3. *If u be an assignment vector, then $\|u\|^2 = \|u\|_0 = 2n$ and*

$$n + \varepsilon^2\alpha m \leq \|\tilde{X}u\|^2 \leq n + 4\varepsilon^2\alpha m,$$

where α is the proportion of clauses in ϕ not satisfied by the true-false assignment corresponding to u . Moreover, if no clause in ϕ contains three true variables, then the lower bound holds with equality.

Proof. To evaluate $\tilde{X}u$, we write

$$\tilde{X}u = \begin{pmatrix} \frac{v_1}{\frac{v_2}{\frac{v_3}{\frac{v_4}{v_5}}}} \end{pmatrix},$$

where $v_1, \dots, v_4 \in \mathbb{R}^n$ and $v_5 \in \mathbb{R}^m$. Clearly $\|v_1\|^2 + \|v_2\|^2 = n$. Since $P\mathbf{1} = 0$ by definition, $v_3 = 0$. Likewise, $v_4 = 0$ because $u_i + u_{i+n} = 1$ for $1 \leq i \leq n$. Write $u^+ \in \mathbb{R}^n$ for the vector consisting of the first n coordinates of u . By definition of Φ , we have

$$\|v_5\|^2 = \varepsilon^2 \|\Phi u^+ - \mathbf{1}\|^2.$$

If the j th clause of ϕ is satisfied by the assignment corresponding to u , then $(\Phi u^+)_j = 1$; otherwise, $1 \leq |(\Phi u^+)_j - 1| \leq 2$. Therefore

$$\alpha m \leq \|\Phi u^+ - \mathbf{1}\|^2 \leq 4\alpha m.$$

If no clause in ϕ contains three true variables, then $|(\Phi u^+)_j - 1| \leq 1$ for all $j \in [m]$, so the lower bound holds with equality. \square

With this choice of \tilde{X} , properties (1) and (3) are easy to establish.

Proposition 4. *If $\text{val}(\phi) = 1$, then there exists a unit vector $u \in \mathbb{R}^{3n}$ such that $\|u\|_0 = 2n$ and*

$$\|\tilde{X}u\|^2 = \frac{1}{2}.$$

Proof. By rescaling, it suffices to produce a vector u such that $\|u\|^2 = \|u\|_0 = 2n$ and

$$\|\tilde{X}u\|^2 = n.$$

Let u be the assignment vector corresponding to a satisfying assignment of ϕ . Applying Proposition 3 yields the claim. \square

Proposition 5. *The matrix \tilde{X} defined in Equation (6) has operator norm at most $3\xi^{-1}$.*

Proof. We employ the following upper bound on the size of the largest singular value due to Schur [17], which is well known (see, e.g., [10]). If r_i is the ℓ_1 norm of the i th row and c_j the ℓ_1 norm of the j th column, then

$$\|\tilde{X}\|_{\text{op}}^2 \leq \max_{i,j} r_i c_j.$$

It is then easy to check that $r_i \leq 3\xi^{-1}$ and $c_j \leq 3\xi^{-1}$. The claim follows. \square

3.3. Proof of Property 2. The remainder of the proof is dedicated to showing that Property 2 holds with $c_1 = 1 + \xi^2$ and c_2 to be specified. For simplicity, we consider vectors u satisfying $\|u\|^2 = 2n$. In what follows, let

$$(7) \quad w \in \underset{u: \|u\|^2=2n, \|u\|_0 \leq 2(1+\xi^2)n}{\text{argmin}} \|\tilde{X}u\|^2.$$

Since w and $-w$ are both minimizers, we assume without loss of generality that w is such that the average value of the last n entries is positive.

We aim to show that, if $\text{val}(\phi) \leq 1 - \alpha$, then

$$(8) \quad \|\tilde{X}w\|^2 \geq (1 + c_2)n$$

for some constant c_2 .

Proposition 3 implies that the value of $\|\tilde{X}u\|$ for an assignment vector is directly related to the number of satisfied clauses in the true-false assignment corresponding to u . To show (8), we will argue that w is “approximately” an assignment vector, so that $\|\tilde{X}w\|$ can still be controlled by $\text{val}(\phi)$.

We interpret $w \in \mathbb{R}^{3n}$ as the concatenation of three vectors w^+ , w^- , and v in \mathbb{R}^n . To show that w is approximately an assignment vector, we need to show that v is close to the all-ones vector, that $w^+ + w^-$ is also close to the all-ones vector, and that w^+ and w^- have almost disjoint support.

Call variable i *good* if exactly one of w_i^+ and w_i^- is zero, and the other lies in the interval $(2/3, 4/3)$. Call clause j *good* if all the variables it contains are good and v_j lies in the interval $(5/6, 7/6)$. Call a clause *bad* if it is not good.

Proposition 6. *Let w be a minimizer in (7). Suppose that there exist positive constants β and γ such that the following two properties hold:*

- $\|w^+\|^2 + \|w^-\|^2 \geq (1 - \beta)n$
- *There are at most γn bad clauses.*

Let

$$\rho = \frac{\varepsilon^2}{36} \left(\frac{9}{13} \alpha - \gamma \right) - \beta.$$

If $\rho > 0$, then then Property 2 holds with $c_1 = 1 + \xi^2$ and $c_2 = \rho$.

Proof. As in the proof of Proposition 3, write

$$\tilde{X}w = \begin{pmatrix} \frac{y_1}{y_2} \\ \frac{y_2}{y_3} \\ \frac{y_3}{y_4} \\ \frac{y_4}{y_5} \end{pmatrix},$$

Then

$$\|y_1\|^2 + \|y_2\|^2 = \|w^+\|^2 + \|w^-\|^2 \geq (1 - \beta)n.$$

Denote by φ the 1-in-3 SAT instance consisting only of the good clauses in ϕ . The vector w induces a true-false assignment to the variables in ϕ in the following way: if the i th variable appears in φ , then it is good, so exactly one of w_i^+ and w_i^- is zero. Set this variable to true if $w_i^+ \neq 0$, and false otherwise. Any assignment to the variables of ϕ must fail to satisfy at least αm clauses, therefore this assignment to the variables of φ must fail to satisfy at least $\alpha m - \gamma n$ clauses.

Suppose that clause j appears in φ and is not satisfied by the true-false assignment corresponding to w . Then clause j contains either 0 true variables or at least 2 true variables. In the former case,

$$(y_5)_j = -\varepsilon v_j < -5\varepsilon/6.$$

In the latter case,

$$(y_5)_j > 4\varepsilon/3 - \varepsilon v_j > \varepsilon/6.$$

We obtain in either case that

$$(y_5)_j^2 > \varepsilon^2/36.$$

Summing over the unsatisfied clauses in φ yields

$$\begin{aligned}\|y_5\|^2 &> (\varepsilon^2/36)(\alpha m - \gamma n) \\ &= \frac{\varepsilon^2}{36} \left(\frac{9}{13}\alpha - \gamma \right) n.\end{aligned}$$

We obtain

$$\begin{aligned}\|\tilde{X}w\|^2 &\geq \|y_1\|^2 + \|y_2\|^2 + \|y_5\|^2 \\ &> (1 - \beta)n + \frac{\varepsilon^2}{36} \left(\frac{9}{13}\alpha - \gamma \right) n \\ &= (1 + \rho)n.\end{aligned}$$

Since w was a minimizer of (7), Property 2 holds with $c_1 = 1 + \xi^2$ and $c_2 = \rho$, as claimed. \square

3.4. Verification of conditions of Proposition 6. In order to verify the conditions of Proposition 6, we require several lemmas about the vector w . Lemma 1 establishes that both v and $w^+ + w^-$ are approximately constant.

Lemma 1. *If $\bar{v} = \frac{1}{n}\mathbf{1}^\top v$, then*

$$\begin{aligned}\sum_{i=1}^n (v_i - \bar{v})^2 &< 2\xi^2 n, \\ \sum_{i=1}^n (w_i^+ + w_i^- - \bar{v})^2 &< 8\xi^2 n, \\ \bar{v}^2 &> 1 - 2\varepsilon^2.\end{aligned}$$

Lemma 2 establishes that for most $i \in [n]$, exactly one of w_i^+ and w_i^- is nonzero.

Lemma 2. *Let*

$$\begin{aligned}I &= \{i : w_i^+ w_i^- \neq 0\} \\ J &= \{j : w_j^+ = 0, w_j^- = 0\}.\end{aligned}$$

If $\varepsilon^2 < 1/4$, then

$$|I| + |J| < 38\xi^2 n.$$

Proofs of both Lemmas appear in the Appendix.

With these Lemmas in hand, we now show that the two conditions of Proposition 6 are satisfied for appropriate choices of β and γ .

Proposition 7. *If $\varepsilon^2 < 1/4$, then*

$$\|w^+\|^2 + \|w^-\|^2 \geq (1 - 45\xi)n.$$

Proof. As in Lemma 2, let

$$\begin{aligned}I &= \{i : w_i^+ w_i^- \neq 0\} \\ J &= \{j : w_j^+ = 0, w_j^- = 0\}.\end{aligned}$$

Let $S = [n] \setminus (I \cup J)$. If $i \in S$, then exactly one of w_i^+ and w_i^- is nonzero.

By Lemma 1,

$$\sum_{i \in S} (w_i^+ + w_i^- - \bar{v})^2 < 8\xi^2 n.$$

Since $\|w\|^2 = 2n$ and $w_i^+ w_i^- = 0$ if $i \in S$, we have the upper bound

$$\sum_{i \in S} (w_i^+ + w_i^-)^2 \leq \|w^+\|^2 + \|w^-\|^2 \leq 2n.$$

On the other hand,

$$\begin{aligned} (w_i^+ + w_i^-)^2 &\geq \bar{v}^2 + (w_i^+ + w_i^- - \bar{v})^2 - 2\bar{v} \cdot |w_i^+ + w_i^- - \bar{v}| \\ &\geq \bar{v}^2 - 2\bar{v} \cdot |w_i^+ - w_i^- - \bar{v}|, \end{aligned}$$

whence the lower bound

$$\begin{aligned} \sum_{i \in S} (w_i^+ + w_i^-)^2 &\geq |S|\bar{v}^2 - 2\bar{v} \sum_{i \in S} |w_i^+ - w_i^- - \bar{v}| \\ &\geq |S|\bar{v}^2 - 4\sqrt{2}\xi n\bar{v}, \end{aligned}$$

where the second inequality follows from the Cauchy-Schwarz inequality and Lemma 1.

By Lemma 2,

$$|S| > n - 38\xi^2 n,$$

hence

$$\begin{aligned} \|w^+\|^2 + \|w^-\|^2 &> n\bar{v}^2 - 38\xi^2 n\bar{v}^2 - 4\sqrt{2}\xi n\bar{v} \\ &> n\bar{v}^2 - 76\xi n - 8\sqrt{2}\xi n > n\bar{v}^2 - 88\xi n, \end{aligned}$$

where we have used the fact that $\xi < \varepsilon < 1/2$.

By orthogonality, Lemma 1, and the fact that $\|w^+\|^2 + \|w^-\|^2 + \|v\|^2 = 2n$, we have

$$n\bar{v}^2 = \|v\|^2 - \sum_{i=1}^n (v_i - \bar{v})^2 > \|v\|^2 - 2\xi^2 n = 2n - (\|w^+\|^2 + \|w^-\|^2) - 2\xi^2 n.$$

Combining the above two displays and rearranging yields

$$\|w^+\|^2 + \|w^-\|^2 > n - 45\xi n,$$

as claimed. \square

Proposition 8. *If $\varepsilon^2 < 1/25$ and $\xi < 1/200$, then there are at most $1056\xi^2 n$ bad clauses.*

Proof. Recall that, by assumption, $\bar{v} \geq 0$. We first show

$$(9) \quad |\bar{v} - 1| < 1/12.$$

Lemma 1 implies

$$\bar{v}^2 > 1 - 2\varepsilon^2 > 23/25,$$

so

$$\bar{v} > \sqrt{23/25} > 11/12.$$

By orthogonality,

$$n\bar{v}^2 \leq \|v\|^2,$$

hence by Proposition 7 and the fact that $\|w^+\|^2 + \|w^-\|^2 + \|v\|^2 = 2n$

$$\bar{v}^2 < 1 + 25\xi < 225/200.$$

Thus

$$\bar{v} < \sqrt{225/200} < 13/12.$$

We now show that most entries of v and $w^+ + w^-$ are near 1. Let

$$L = \{\ell : |v_\ell - 1| > 1/6\},$$

$$K = \{k : |w_k^+ + w_k^- - 1| > 1/3\}.$$

If $\ell \in L$, then by (9),

$$(v_\ell - \bar{v})^2 > 1/36.$$

Likewise, if $k \in K$, then

$$(w_k^+ + w_k^- - \bar{v})^2 > 1/16$$

Applying Lemma 1 yields

$$(10) \quad |L| < 36 \sum_{i=1}^n (v_i - \bar{v})^2 < 72\xi^2 n,$$

$$(11) \quad |K| < 16 \sum_{i=1}^n (w_i^+ + w_i^- - \bar{v})^2 < 128\xi^2 n.$$

By Lemma 2 and (11), there exists a set $G \in [n]$ of size at least

$$(1 - 164\xi^2)n$$

such that for all $i \in G$:

- Exactly one of w_i^+ and w_i^- is nonzero, and
- $|w_i^+ + w_i^- - 1| < 1/3$.

By definition, all variables in G are good.

Since ϕ is 6-bounded, the other $164\xi^2 n$ variables are contained in at most $984\xi^2 n$ clauses. This fact combined with (10) implies that there are at most

$$(984\xi^2 + 72\xi^2)n = 1056\xi^2 n$$

bad clauses, as claimed. \square

We are now ready to prove Theorem 1 for a specific choice of δ .

Proposition 9. *Let $\varepsilon = \sqrt{1/50}$ and $\xi \in (0, 1/200)$ be chosen small enough that*

$$\rho = \frac{1}{7200} \left(\frac{9}{13} \alpha - 1056\xi^2 \right) - 45\xi > 0.$$

If \tilde{X} is defined as in Section 3.2, then \tilde{X} satisfies the three properties given in Section 3.1, with $c_1 = 1 + \xi^2$, $c_2 = \rho$, and $c_3 = 9\xi^{-1}$. Therefore Theorem 1 holds with

$$\delta = 1 - \frac{1 + \rho}{18\xi^{-1}}, \quad \lambda_1 = c_1, \quad \lambda_2 = \frac{18\xi^{-1}}{18\xi^{-1} - \rho}.$$

Proof. Properties 1 and 3 have been shown to hold in Propositions 4 and 5. By Propositions 7 and 8, if $\varepsilon^2 < 1/25$ and $\xi < 1/200$, then the conditions of Proposition 6 hold with $\beta = 45\xi$ and $\gamma = 1056\xi^2$.

$$\beta = 45\xi$$

$$\gamma = 1056\xi^2.$$

By assumption, $\rho > 0$, so Property 2 holds with $c_1 = 1 + \xi^2$ and $c_2 = \rho$. \square

3.5. Extension to general δ . Finally, Theorem 1 follows from the following Proposition.

Proposition 10. *If Theorem 1 holds for some $\delta \in (0, 1)$, then it holds for any $\delta' \in (0, 1)$.*

Proof. We first show the claim for $\delta' \in (0, \delta)$.

Given a matrix $X \in \mathbb{R}^{n \times p}$, set

$$M = \frac{1}{1+c} (X^\top X + cI_{p \times p}) ,$$

where $I_{p \times p}$ is the identity matrix and $c > 0$ is a suitable constant to be chosen later. The matrix $cI_{p \times p}$ is clearly positive semidefinite and the matrix $X^\top X$ is a Gram matrix and hence positive semidefinite as well. Therefore M is also positive semidefinite and possesses a factorization $M = \hat{X}^\top \hat{X}$.

Then $X \in \text{RIP}(k, \delta)$ if and only if $\hat{X} \in \text{RIP}(k, \delta/(1+c))$. Indeed, the construction of \hat{X} implies

$$\|\hat{X}u\|^2 = u^\top M u = \frac{1}{1+c} (\|Xu\|^2 + c\|u\|^2) .$$

In particular, $\|Xu\|^2 = (1+\theta)\|u\|^2$ for some $\theta \in \mathbb{R}$, then $\|\hat{X}u\|^2 = \left(1 + \frac{\theta}{1+c}\right) \|u\|^2$.

Set $c = \frac{\delta}{\delta'} - 1 > 0$. Then deciding whether $X \in \text{RIP}(k, \delta)$ or $X \notin \text{RIP}(k/\lambda_1, \lambda_2\delta)$ is reducible in polynomial time to the problem of deciding whether $\hat{X} \in \text{RIP}(k, \delta')$ or $\hat{X} \notin \text{RIP}(k/\lambda_1, \lambda_2\delta')$. If the former problem is NP-hard, then the latter is as well.

On the other hand, for $\delta' \in (\delta, 1)$, we recall that we can assume $\|Xu\| \leq \|u\|$ for all u . Given such a matrix X , let $\hat{X} = \sqrt{\frac{1-\delta'}{1-\delta}} X$. If $X \in \text{RIP}(k, \delta)$ and $\|Xu\| \leq \|u\|$ for all u , then $\hat{X} \in \text{RIP}(k, \delta')$ and $\|\hat{X}u\| \leq \|u\|$ for all u . If $X \notin \text{RIP}(k/\lambda_1, \lambda_2\delta)$ and $\|Xu\| \leq \|u\|$ for all u , then $\hat{X} \notin \text{RIP}(k/\lambda_1, \lambda_2'\delta')$ for $\lambda_2' > 1$ such that $\lambda_2'\delta' = 1 - \frac{1-\lambda_2\delta}{1-\delta}(1-\delta')$. Hence deciding whether $X \in \text{RIP}(k, \delta)$ or $X \notin \text{RIP}(k/\lambda_1, \lambda_2\delta)$ is reducible in polynomial time to deciding whether $\hat{X} \in \text{RIP}(k, \delta')$ or $\hat{X} \notin \text{RIP}(k/\lambda_1, \lambda_2'\delta')$, and the latter problem is NP-hard is the former one is. \square

4. CONCLUSION

In this work, we show that it is NP-hard to certify the Restricted Isometry Property, even approximately, for all $\delta \in (0, 1)$. This resolves a question implicit in earlier work, which either required $\delta = 1 - o(1)$ or relied on stronger assumptions than $P \neq NP$.

Proposition 1 is of independent interest—though similar gap hardness results exist in the literature, the bounded variant may be of use in other reductions.

We note that we have made no attempt to optimize the constants in the proof of Theorem 1, but even a more careful version of this proof will still produce λ_1 and λ_2 very close to 1. It is an open question whether NP-hardness can be proven for a version of Problem 2 in which the constants λ_1 and λ_2 are large.

The most important open question in this area is to show that certifying $X \in \text{RIP}(k, \delta)$ is hard *on average* when X is drawn from some natural probability distribution. In particular, as noted above, there are many random constructions known to generate RIP matrices with high probability [4, 11]. Our work does *not* rule out a computationally efficient procedure which is able to certify $X \in \text{RIP}(k, \delta)$ when

X is generated randomly. A proof of the (non)existence of such a procedure would be a very important theoretical result.

REFERENCES

1. Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy, *Proof verification and the hardness of approximation problems*, J. ACM **45** (1998), no. 3, 501–555, Preliminary version in FOCS '92. MR 1639346
2. Afonso S. Bandeira, Edgar Dobriban, Dustin G. Mixon, and William F. Sawin, *Certifying the restricted isometry property is hard*, IEEE Trans. Inform. Theory **59** (2013), no. 6, 3448–3450. MR 3061257
3. Afonso S. Bandeira, Matthew Fickus, Dustin G. Mixon, and Percy Wong, *The road to deterministic matrices with the restricted isometry property*, J. Fourier Anal. Appl. **19** (2013), no. 6, 1123–1149. MR 3132908
4. Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin, *A simple proof of the restricted isometry property for random matrices*, Constr. Approx. **28** (2008), no. 3, 253–263. MR 2453366
5. Jean Bourgain, Stephen Dilworth, Kevin Ford, Sergei Konyagin, and Denka Kutzarova, *Explicit constructions of RIP matrices and related problems*, Duke Math. J. **159** (2011), no. 1, 145–185. MR 2817651
6. Emmanuel J. Candès, *The restricted isometry property and its implications for compressed sensing*, C. R. Math. Acad. Sci. Paris **346** (2008), no. 9–10, 589–592. MR 2412803
7. Emmanuel J. Candès and Terence Tao, *Decoding by linear programming*, IEEE Trans. Inform. Theory **51** (2005), no. 12, 4203–4215. MR 2243152
8. Ronald A. DeVore, *Deterministic constructions of compressed sensing matrices*, J. Complexity **23** (2007), no. 4–6, 918–925. MR 2371999
9. Uriel Feige, *A threshold of $\ln n$ for approximating set cover*, J. ACM **45** (1998), no. 4, 634–652.
10. Gene H Golub and Charles F Van Loan, *Matrix computations*, fourth ed., JHU Press, 2013.
11. Ishay Haviv and Oded Regev, *The restricted isometry property of subsampled fourier matrices*, Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10–12, 2016, 2016, pp. 288–297.
12. Sanjeev Khanna and Madhu Sudan, *The optimization complexity of constraint satisfaction problems*, Electronic Colloquium on Computational Complexity (ECCC) **3** (1996), no. 28.
13. Pascal Koiran and Anastasios Zouzias, *Hidden cliques and the certification of the restricted isometry property*, IEEE Trans. Inform. Theory **60** (2014), no. 8, 4999–5006. MR 3245368
14. Abhiram Natarajan and Yi Wu, *Computational complexity of certifying restricted isometry property*, Approximation, randomization, and combinatorial optimization, LIPIcs. Leibniz Int. Proc. Inform., vol. 28, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2014, pp. 371–380. MR 3319003
15. Christos H. Papadimitriou and Mihalis Yannakakis, *Optimization, approximation, and complexity classes*, J. Comput. Syst. Sci. **43** (1991), no. 3, 425–440, Preliminary version in STOC '88.
16. Thomas J. Schaefer, *The complexity of satisfiability problems*, Proceedings of the 10th Annual ACM Symposium on Theory of Computing, May 1–3, 1978, San Diego, California, USA, 1978, pp. 216–226.
17. J. Schur, *Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen*, J. Reine Angew. Math. **140** (1911), 1–28. MR 1580823
18. Andreas M. Tillmann and Marc E. Pfetsch, *The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing*, IEEE Trans. Inform. Theory **60** (2014), no. 2, 1248–1259. MR 3164973
19. Vijay V Vazirani, *Approximation algorithms*, Springer Science & Business Media, 2013.

APPENDIX A. PROOFS OF LEMMAS 1 AND 2

Lemma 1. *If $\bar{v} = \frac{1}{n} \mathbf{1}^\top v$, then*

$$(12) \quad \sum_{i=1}^n (v_i - \bar{v})^2 < 2\xi^2 n,$$

$$(13) \quad \sum_{i=1}^n (w_i^+ + w_i^- - \bar{v})^2 < 8\xi^2 n,$$

$$(14) \quad \bar{v}^2 > 1 - 2\varepsilon^2.$$

Proof. We first show a simple upper bound on $\|\tilde{X}w\|^2$. Let $u = (0, \dots, 0, 1, \dots, 1)$ be the assignment vector corresponding to the assignment that sets each variable to false. By Proposition 3,

$$\|\tilde{X}u\|^2 = n + \varepsilon^2 m \leq (1 + \varepsilon^2)n.$$

Since w satisfies (7),

$$(15) \quad \|\tilde{X}w\| \leq \|\tilde{X}u\| \leq (1 + \varepsilon^2)n < 2n,$$

where we have used the assumption that $\varepsilon < 1$.

By definition,

$$\|Pv\|^2 = \|(I - \frac{1}{n} \mathbf{1} \mathbf{1}^\top)v\|^2 = \sum_{i=1}^n (v_i - \bar{v})^2.$$

By (15),

$$\|\xi^{-1}Pv\|^2 \leq \|\tilde{X}w\|^2 < 2n,$$

hence

$$\sum_{i=1}^n (v_i - \bar{v})^2 < 2\xi^2 n,$$

as claimed.

For the second bound, by Young's inequality,

$$\sum_{i=1}^n (w_i^+ + w_i^- - \bar{v})^2 \leq 2 \sum_{i=1}^n (w_i^+ - w_i^- - v_i)^2 + 2 \sum_{i=1}^n (v_i - \bar{v})^2.$$

Note that

$$\begin{aligned} \xi^{-2} \sum_{i=1}^n (w_i^+ - w_i^- - v_i)^2 &= \|\xi^{-1}Iw^+ + \xi^{-1}Iw^- - \xi^{-1}Iv\|^2 \\ &\leq \|\tilde{X}w\|^2. \end{aligned}$$

By (15), this quantity is smaller than $2n$. Combining this with (12) and (13) yields

$$2 \sum_{i=1}^n (w_i^+ - w_i^- - v_i)^2 + 2 \sum_{i=1}^n (v_i - \bar{v})^2 < 8\xi^2 n.$$

For the third inequality, by (15),

$$\|w^+\|^2 + \|w^-\|^2 \leq \|\tilde{X}w\|^2 \leq (1 + \varepsilon^2)n.$$

Therefore

$$\|v\|^2 = 2n - (\|w^+\|^2 + \|w^-\|^2) \geq (1 - \varepsilon^2)n.$$

By orthogonality and (12),

$$\|v\|^2 = \|v - \bar{v}\mathbf{1}\|^2 + \|\bar{v}\mathbf{1}\|^2 = \sum_{i=1}^n (v_i - \bar{v})^2 + n\bar{v}^2 < 2\xi^2 n + n\bar{v}^2.$$

We obtain

$$n\bar{v}^2 > \|v\|^2 - 2\xi^2 n > (1 - \varepsilon^2)n - \varepsilon^2 n = (1 - 2\varepsilon^2)n,$$

and the claim follows. \square

Lemma 2. *Let*

$$\begin{aligned} I &= \{i : w_i^+ w_i^- \neq 0\} \\ J &= \{j : w_j^+ = 0, w_j^- = 0\}. \end{aligned}$$

If $\varepsilon^2 < 1/4$, then

$$|I| + |J| < 38\xi^2 n.$$

Proof. If w_j^+ and w_j^- are both 0 for some j , then

$$(w_j^+ + w_j^- - \bar{v})^2 = \bar{v}^2 > (1 - 2\varepsilon^2) > 1/2$$

by Lemma 1. Hence

$$|J| \leq 2 \sum_{i=1}^n (w_i^+ + w_i^- - \bar{v})^2 < 16\xi^2 n.$$

We now show that v has almost full support. By Lemma 1,

$$\sum_{i=1}^n (v_i - \bar{v})^2 < 2\xi^2 n.$$

If $v_i = 0$, then by Lemma 1

$$(v_i - \bar{v})^2 = \bar{v}^2 > (1 - 2\varepsilon^2) > 1/2.$$

If p is the number of zero entries in v , then

$$p \leq 2 \sum_{i=1}^n (v_i - \bar{v})^2 < 4\xi^2 n,$$

Therefore

$$\|v\|_0 = n - p > (1 - 4\xi^2)n.$$

Since $\|w\|_0 \leq 2c_1 n = 2n + 2\xi^2 n$, we have

$$\|w^+\|_0 + \|w^-\|_0 = \|w\|_0 - \|v\|_0 < (1 + 6\xi^2)n.$$

We obtain

$$\begin{aligned} |I| &= \|w^+\|_0 + \|w^-\|_0 + |J| - n \\ &< (1 + 6\xi^2)n + 16\xi^2 n - n = 22\xi^2 n. \end{aligned}$$

Combining the above bounds on $|I|$ and $|J|$ yields the claim. \square

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 77 MASSACHUSETTS AVENUE, CAMBRIDGE, MA 02139, USA

E-mail address: jweed@mit.edu